

## 短期記憶を用いたシンプルリカレントネットワークによるカオス 時系列の短期予測

菊地 進一 中西正和

慶應義塾大学大学院 理工学研究科 開放環境科学専攻

### 概要

時系列の非線形予測手法の1つとして Elman のリカレントニューラルネットワークがある。筆者らは、この構造の入力に過去の履歴を直接与えることで、学習が高速化されることを確認している。ただし、これまでは周期的なデータに対する適用であったため、過去の履歴が重要であった。本研究は、カオス的なデータに対しても、提案手法が有効であることを示すものである。この構造にクロスエントロピー法を用いることにより、従来の Elman の手法と比較して、汎化能力を低下させずに、学習が数倍から 130 倍程度まで高速化されることを確認した。

## Short-term Forecasting of Chaotic Time Series by Simple Recurrent Network with Short-term Memory

Shimichi Kikuchi Masakazu Nakanishi

School of Science for Open and Environmental Systems, Keio University

### Abstract

Recently, we have proposed a structure which uses directly past histories in Elman's recurrent neural network. It is shown that learning is accelerated using the structure. Data in the previous work is periodic, so past histories are important for learning time series. In this work, it is shown that the structure is also useful for chaotic data. Moreover, we employ the cross entropy method for fast learning. Using our method, it is shown that learning speed attains 130 times faster without making generalization ability worse compared with Elman's method.

### 1 はじめに

決定論的カオス時系列の短期予測という概念は、複雑な様相を呈する実世界の時系列に対しても、工学的アプローチが有効であるという新たなカオス工学の分野を開いた。カオス時系列の短期予測の手法としては、局所的に多項式をあてはめる方法 [1]、局所動径基底関数を用いる方法 [2]、シンプレクス投影法 [3]、テセレーション法 [4]、ウェーブレットを用いる方法 [5] など種々の方法が提案されている。

中でも、ニューラルネットワークは時系列の非線形予測手法としても、最もよく研究されている分野の1つであり [6]、そこで得られた新たな見解がカオス時系列に対してどのように働くのかを調べるのは非常に重要なことである。しかし、従来のニューラルネットワークを用いた時系列の研究ではカオス時系列を意識した研究というものは少なかった。また、カオス工学の研究でも、ニューラルネットワークの研究で得られた最新の知見を応用していることが少なかった。

そこで、本研究では、最近、筆者らの提案したリカレントニューラルネットワークである短期記憶を用いたシンプルリカレントネットワーク [7][8] がカオス時系列の学習に対しても有効な手法であることを示す。これは Elman のネットワーク [9] の入力に直接過去の履歴を与えるネットワークである。このネットワークを用いることにより、汎化能力を低下させずに、Elman の手法と比較して、学習時間が短くなる、収束率が向上するなどの有効性が示されている [7][8]。

本研究では学習アルゴリズムとして、誤差逆伝搬学習の高速化手法の1つであるクロスエントロピー法 [10] を用いる。クロスエントロピー法は、従来の静的な問題に適用されていた手法であり、時系列の学習の問題において、その高速化能力や汎化能力はあまり示されていなかった。本論文では、クロスエントロピー法が、カオス的な時系列の学習においても汎化能力を低下させない有効な高速化手法であることを示す。

以下、本論文では、カオス時系列の学習に際し、これらを用いると、Elmanの手法と比較して、汎化能力を低下させずに、学習速度を数倍から130倍程度まで高速化することが可能であることを示す。

## 2 短期記憶を用いたシンプルリカレントネットワーク

筆者らは、入力にバッファを持たせた図1のような短期記憶を用いたシンプルリカレントネットワーク (Simple Recurrent Network with Short-term Memory, SRNSM) を提案した [7][8]。この構造により、時間的に近いデータに対する学習が明確になり、学習速度が向上する。実際に、非常にシンプルな改良であるにも関わらず、周期的な時系列に対して、汎化能力を低下させずに、学習が数十倍も高速化されることが示されている [7][8]。

## 3 クロスエントロピー法

本研究では、学習速度と収束率の向上のために、クロスエントロピー法 [10] を用いる。クロスエントロピー法では、誤差評価関数として、通常平均出力自乗誤差項ではなく、次の評価関数  $C$  を用いる。

$$C = - \sum_i \{ t_i \ln o_i + (1 - t_i) \ln(1 - o_i) \} \quad (1)$$

ここで、 $o_i$  は出力ユニット  $i$  の出力値、 $t_i$  はその目標出力値である。リンク重みの更新には、この評価関数  $C$  をリンク重みで偏微分した値を用いる。

## 4 ノイズを付加された Hénon 写像列

以下、4章と5章では、コンピュータシミュレーションによる実験結果について述べる。4章を実験1、5章を実験2と呼ぶことにする。問題の設定は、ある時刻  $t$  の値を入力とし、次の時刻  $t+1$  の値を出力として学習するものとする。SRNSMでは、入力としてバッファを持たせるので、バッファの長さを  $l$  とすると、入力に  $x(t), x(t-1), x(t-2), \dots, x(t-l+1)$  を与えればよい。

本章では、簡単なダイナミクスを持つカオス時系列の例として、次式の Hénon 写像 [11] により生成された時系列  $x(t)$  にノイズ  $\varepsilon(t)$  を付加した時系列  $x'(t)$  を取り扱う。

$$x(t) = 1 - ax(t-1)^2 + bx(t-2) \quad (2)$$

$$x'(t) = x(t) + \varepsilon(t) \quad (3)$$

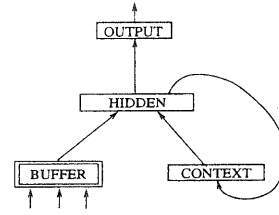


図 1: 短期記憶を用いたシンプルリカレントネットワーク

本研究では、 $a = 1.4, b = 0.3, x(0) = 0.046, x(1) = -1.04$  を用いた。ノイズには観測誤差として知られる正規乱数を用いた。 $\varepsilon(t)$  は平均 0、分散  $0.1^2$  の正規乱数列である。本実験では、学習用データに 200 点、テスト用データに 200 点を用いた。

表 1 に、バッファの長さを変化させた時の誤差逆伝搬学習 (BP) とクロスエントロピー法 (CE) の学習回数を示す。また、汎化能力の基準として、0 から 1 に正規化されたテストデータに対する平均出力自乗誤差 (予測誤差と呼ぶ) とテストデータに対する実測値と予測値の間の相関係数の 2 つの値を示す。ここで、Buffer=1 で BP の欄が、従来の Elman の手法である。各パラメータの値は、学習係数  $\eta = 0.01$ 、慣性係数  $\alpha = 0.9$  とした。学習回数は平均出力自乗誤差が 0.0040 になるまでの 100 回の試行の平均エポック数である。ここでは、200 回を 1 エポックとした。収束率はいずれも 100% である。隠れユニット数については、リンク重みに学習内容が記憶されることを考慮し、モデルやアルゴリズムではなくリンク数で差が生じないように、リンク数がほぼ同じになるように設定した。Elman のネットワークの隠れユニット数が 30 個 (リンク数 960 本) である。また、図 2 に、3 つの学習曲線を示す。“Elman” は Elman の手法の結果、“SRNSM” は Buffer=4 のものに 2 つのアルゴリズムを適用した結果である。

表 1 より、Buffer の概念を導入すると、汎化能力を低下させずに、学習が高速化されることが分かる。また、クロスエントロピー法は、カオス時系列の学習においても学習を高速化する能力があることが分かる。さらに、SRNSM との相性も良い。具体的には、Buffer=4 の SRNSM に CE を適用した本手法の結果と Elman の手法の結果を比較すると、学習速度は 6.8 倍高速になった。図 2 を見ても、SRNSM と CE による学習が振動などしない安定した高速化手法であることが分かる。図 3 に、Buffer=4 の SRNSM に

表 1: 実験 1 の結果

l	学習回数		予測誤差		相関係数	
	BP	CE	BP	CE	BP	CE
1	8524	3428	0.0029	0.0031	0.971	0.968
2	7658	2707	0.0027	0.0029	0.973	0.971
3	6745	2536	0.0030	0.0031	0.971	0.969
4	6445	1246	0.0035	0.0033	0.967	0.967
5	6417	2727	0.0036	0.0039	0.967	0.965

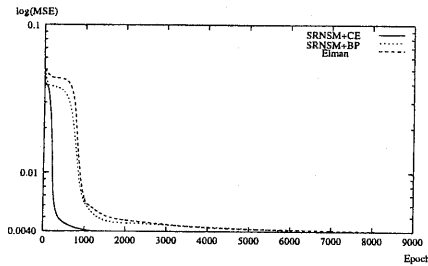


図 2: 実験 1 における学習曲線

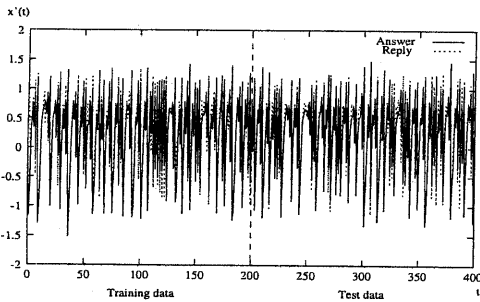


図 3: 実験 1 における学習結果と予測結果

CE を適用した時の学習結果と予測結果をグラフにしたものを示す。実線が正解で、破線がネットワークの出力である。このグラフからも、本手法によるカオス時系列に対する予測精度の高さが分かる。

実験 1 の結果について考察する。本実験では最適なバッファ数が 4 になっている。これは過去の時間を 2 時刻までさかのぼる元のダイナミクスに非常に近い値である。これより、SRNSM は直接的に過去の値を入力することで、ダイナミクスの学習を簡単にしていることが分かる。また、ノイズが付加されたせいで、SRNSM がダイナミクスを学習するのに、より多くの過去の情報を必要としていることが分かる。すなわち、最適なバッファ数のみから隠れたダイナミクスの形を簡単に決定するのは危険である。また、SRNSM はバッファ数さえあれば、ノイズがあってもダイナミクスを学習できる、ロバスト性の

高い手法であることが分かる。

## 5 はしかの患者数

本章では、現実的なカオス時系列データの例として、1928 年 1 月から 1972 年 6 月までの New York City の毎月のはしかの患者数のデータ [12] を取り扱う。このデータに関する研究としては、文献 [3][13][14] などがある。本実験では、1928 年 1 月から 1947 年 12 月までの値を学習用データ (240 点) に、1948 年 1 月から 1972 年 6 月までの値を予測用データ (294 点) に用いた。

表 2 に、バッファの長さを変化させた時の学習回数、予測誤差、相関係数を示す。ここで、Buffer=1 で BP の欄が、従来の Elman の手法である。各パラメータの値は、学習係数  $\eta = 0.01$ 、慣性係数  $\alpha = 0.9$  とした。学習回数は平均出力自乗誤差が 0.0030 になるまでの 100 回の試行の平均エポック数である。ここでは、240 回を 1 エポックとした。収束率はいずれも 100% である。隠れユニット数の設定については、実験 1 と同様である。また、図 4 に、3 つの学習曲線を示す。“Elman” は Elman の手法の結果、“SRNSM” は Buffer=11 のものに 2 つのアルゴリズムを適用した結果である。

表 2 より、Buffer の概念を導入すると、汎化能力を低下させずに、学習が高速化されることが分かる。また、クロスエントロピー法は、現実的なカオス時系列の学習においても学習を高速化することが分かる。さらに、SRNSM との相性も良い。具体的には、Buffer=11 の SRNSM に CE を適用した本手法の結果と Elman の手法の結果を比較すると、学習速度は 128 倍も高速になった。図 4 を見ても、SRNSM と CE による学習が振動などしない安定した高速化手法であることが分かる。図 5 に、Buffer=11 の SRNSM に CE を適用した時の学習結果と予測結果をグラフにしたものを示す。実線が正解で、破線がネットワークの出力である。このグラフからも、本手法によるカオス時系列に対する予測精度の高さが分かる。

実験 2 の結果について考察する。本実験では最適なバッファ数が 11 になっている。この長さは、実験 1 の結果より、ダイナミクスが難しい点と観測ノイズの両方によるものと考えられる。また、学習が非常に難しいため、SRNSM の高速化能力が非常に機能している。このことから、SRNSM は現実的なカオス時系列の学習に極めて有効であることが期待できる。特に、128 倍も学習が高速化されているのにも関わらず、汎化能力が低下していない点は非常に

表 2: 実験 2 の結果

l	学習回数		予測誤差		相関係数	
	BP	CE	BP	CE	BP	CE
1	16992	4676	0.0029	0.0029	0.961	0.961
3	6446	1656	0.0029	0.0028	0.966	0.968
5	5968	1372	0.0031	0.0030	0.966	0.968
7	2040	437	0.0031	0.0031	0.966	0.965
9	864	166	0.0029	0.0029	0.971	0.971
11	641	133	0.0028	0.0029	0.971	0.970
13	530	113	0.0027	0.0031	0.965	0.963
15	520	104	0.0031	0.0036	0.963	0.955

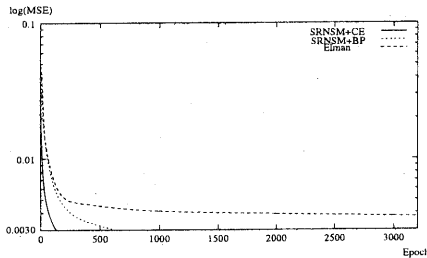


図 4: 実験 2 における学習曲線

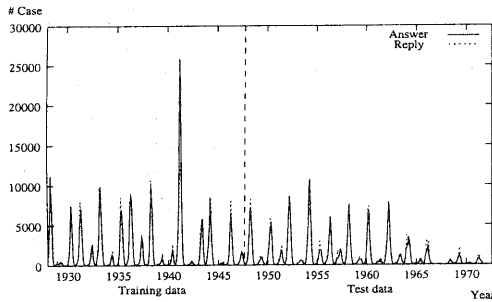


図 5: 実験 2 における学習結果と予測結果

有用な性質であると言える。

## 6 おわりに

カオス時系列の学習に際し、構造として SRNSM、学習アルゴリズムとして CE を用いることにより、汎化能力を低下させずに学習速度が数倍から 130 倍程度まで向上することが示された。また、本手法は観測ノイズに強いことも示された。特に、現実的なカオス時系列であるはしかの患者数のデータにおいては、従来の Elman の手法と比較して、汎化能力を低下させずに 128 倍も学習が高速化されるという結果が得られた。これらは、ニューラルネットワークを用いたカオス時系列の学習に際して、非常に重要な性質であると考えられる。

今後の課題としては、最適なバッファ数とダイナミクスとの関係に対する定性的な考察が残されている。これに関しては、フラクタル次元 [15] などによる理論的な解析が有効であると考えている。

## 参考文献

- [1] J.D. Farmer and J.J. Sidorowich: Predicting Chaotic Time Series, *Physical Review Letters*, Vol. 59, No. 8, pp. 845-848 (1987).
- [2] M. Casdagali: Nonlinear Prediction of Chaotic Time Series, *Physica D*, Vol. 35, No. 3, pp.335-356 (1989).
- [3] G. Sugihara and R.M. May: Nonlinear Forecasting as a Way of Distinguishing Chaos from Measurement Error in Time Series, *Nature*, Vol. 344, pp. 734-741 (1990).
- [4] A.I. Mees: Dynamical Systems and Tesselations: Detecting Determinism in Data, *International Journal of Bifurcation and Chaos*, Vol. 1, No. 4, pp. 777-794 (1991).
- [5] L. Cao, Y. Hong, H. Fang and G. He: Predicting Chaotic Time Series with Wavelet Networks, *Physica D*, Vol. 85, No. 1&2, pp. 225-238 (1995).
- [6] B.A. Pearlmutter: Gradient Calculations for Dynamic Recurrent Neural Networks: A Survey, *IEEE Trans. on Neural Networks*, Vol. 6, No. 5, pp. 1212-1228 (1995).
- [7] 菊地進一, 中西正和: 短期記憶を用いたリカレントニューラルネットワークと高速な構造学習法, 信学技報, NC99-67 (1999).
- [8] S. Kikuchi and M. Nakanishi: Learning Time Series by Simple Recurrent Network with Short-term Memory, *Proc. of IASTED MIC*, now printing (2000).
- [9] J.L. Elman: Finding Structure in Time, *Cognitive Science*, Vol. 14, pp. 179-211 (1990).
- [10] A. van Ooyen and B. Nienhuis: Improving the Convergence of the Back-propagation Algorithm, *Neural Networks*, Vol. 5, No. 3, pp. 465-471 (1992).
- [11] M. Hénon: A Two-dimensional Mapping with a Strange Attractor, *Communications in Mathematical Physics*, Vol. 50, pp. 69-77 (1986).
- [12] J.A. Yorke and W.P. London: Recurrent Outbreaks of Measles, Chickenpox and Mumps II. Systematic Differences in Contact Rates and Stochastic Effects, *American Journal of Epidemiology*, Vol. 98, No. 6, pp. 469-482 (1973).
- [13] W.M. Schaffer and M. Kot: Nearly One Dimensional Dynamics in an Epidemic, *Journal of Theoretical Biology*, Vol. 112, pp. 403-427 (1985).
- [14] L.F. Olsen and W.M. Schaffer: Chaos Versus Noisy Periodicity: Alternative Hypotheses for Childhood Epidemics, *Science*, Vol. 249, pp. 499-504 (1990).
- [15] T. Higuchi: Approach to an Irregular Time Series on the Basis of the Fractal Theory, *Physica D*, Vol. 31, pp. 277-283 (1988).